APEX VERIFY

APEX VERIFY RESEARCH / WHITE PAPER NO. 02

Proof at the point of capture.

Synthetic media, C2PA and capture-first provenance for U.S. real estate and insurance.

URBAN HERAK APEX VERIFY OCTOBER 29, 2026 64 SOURCES

ABSTRACT

U.S. real estate brokerages and insurers should stop trying to decide after the fact whether a photo or video "looks real". Instead, they should require media whose origin was recorded at the moment of capture, and treat everything else as unverified. The losses are measurable. The FBI logged $275.1 million in real estate fraud losses in 2025, up 58.5% in a year [1]. 59% of title firms faced a seller-impersonation attempt in the past year, and 58% have encountered deepfaked images or voices [2]. In claims, 99% of surveyed insurers have seen manipulated or AI-altered documentation and 36% of consumers would consider altering a claim photo [3].

Neither common defense holds. EXIF metadata can be rewritten by anyone who holds the file and is routinely stripped when files are shared [4]. Detectors lose much of their accuracy on generators they never saw, a pixel-level attack can defeat them almost completely, and human reviewers perform near chance [5, 6, 7]. C2PA Content Credentials provide the right cryptographic design, a signed manifest bound to the exact bytes. But hardware-backed capture signing is shipping at scale on essentially one phone line, and embedded credentials get stripped too [8, 9].

A capture-first model fits the regulations that are now arriving. It pairs a capture-only camera with an off-file integrity record, a permanent public verification link for media meant to be published, and a revocable 24-hour private link for sensitive evidence. California's AB 723 already requires a public link to the "original, unaltered image" behind every altered listing photo [10]. Apex Verify follows this model. Its claim is narrow: this file came from an Apex Camera capture and is unchanged since. It is not an AI detector. It does not prove what a scene depicts or that it was not staged. It does not yet embed certificate-backed, interoperable C2PA credentials in every export, so its records cannot yet be read by standard C2PA validators.

CONTENTS

01 Real estate: $275 million in losses, and photos lend the credibility

02 Insurance: a $308.6 billion baseline as policyholders take the pictures

03 Why metadata fails

04 Why detection fails

05 C2PA: the right cryptography, not yet the coverage

06 Regulation and evidence: converging on provenance without mandating a technology

07 Capture-first provenance and two link types

08 Honest limits

09 Conclusion

A Key answers

– References

01

Real estate: $275 million in losses, and photos lend the credibility

1.1

Real estate fraud is mostly about identity

The FBI's Internet Crime Complaint Center counts 12,368 real estate complaints and $275,110,419 in losses for 2025. The comparable figures are $173.6 million for 2024 and $145.2 million for 2023, so losses rose 89% in two years [1]. Victims aged 60 and over account for about $123.7 million of the 2025 total, roughly 45%.

$275.1M

real estate fraud losses reported to the FBI IC3 in 2025, up 58.5% in a year [1]

89%

rise in reported real estate losses in two years, from $145.2 million in 2023 [1]

59%

of title firms reported a seller-impersonation attempt in the prior year, up from 28% in 2024 [12, 2]

58%

of title professionals had encountered deepfaked images or voices [2]

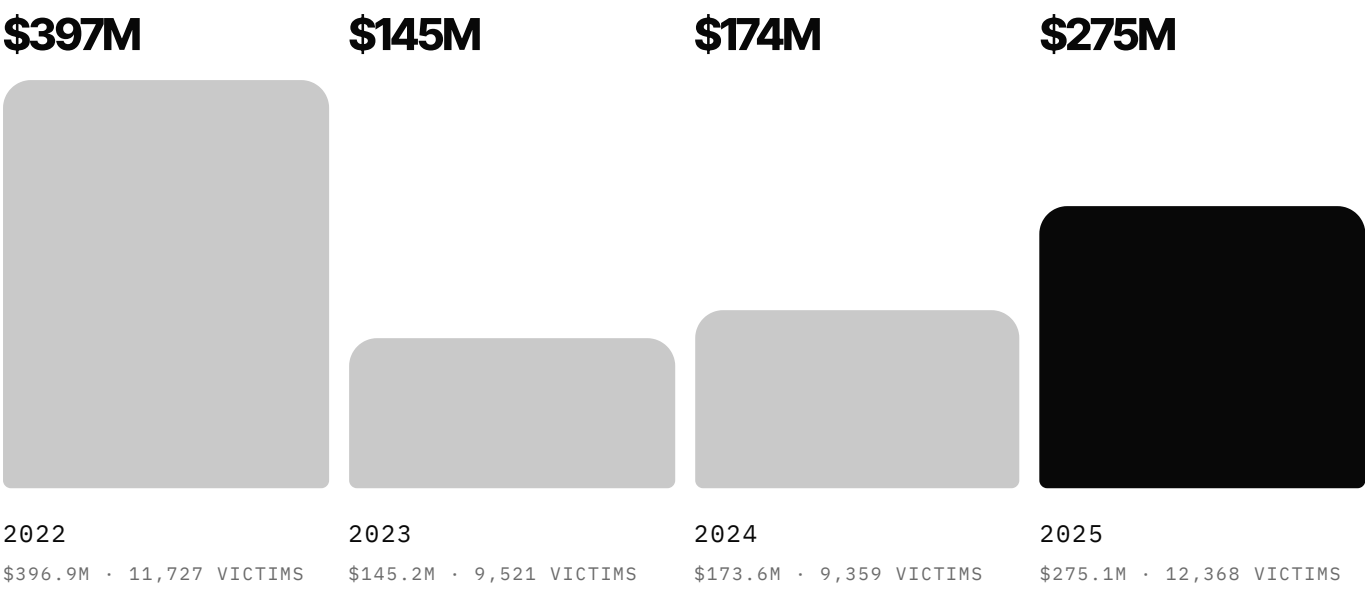

Figure 1. Real estate fraud losses and victim counts reported to the FBI Internet Crime Complaint Center, 2022–2025. Losses fell from $396.9 million in 2022 to $145.2 million in 2023, then rose to $173.6 million in 2024 and $275.1 million in 2025 [1, 64].

The real estate count understates the problem, because IC3 assigns each complaint a single crime type. Closing-wire fraud often lands in business email compromise (BEC), which cost $3.05 billion in 2025. IC3's BEC definition explicitly covers compromised "virtual meeting applications", and it names voice cloning as a way to request wire payments [1].

In 2025 IC3 also reported, for the first time, an "AI Related" descriptor: 22,364 complaints and $893.3 million in adjusted losses. Only 115 real estate complaints ($2.7 million) carried it. The descriptor applies only when the victim mentions AI, and there is no earlier year to compare against. IC3 itself warns that victims often "do not realize the extent AI may be involved", so the small figure is a floor, not evidence that AI plays a minor role [1].

Rental fraud shows how real photos serve fake identities. The FTC counts nearly 65,000 rental-scam reports and about $65 million in losses from January 2020 to June 2025, with a $1,000 median loss. Its main technique is copying a legitimate listing, often a property actually for sale, changing the contact details and reposting it. Only 4.8% of mass-market fraud victims complain to a government body or the BBB, so reported figures understate actual harm [11].

Sales fraud follows the same pattern at higher stakes. The American Land Title Association's 2026 study surveyed 245 title professionals. 59% reported a seller-impersonation attempt in the prior year, up from 28% in 2024, and 45% in the prior 30 days. Vacant land was the top target, cited by 82% [12, 2]. 58% had encountered deepfaked images or voices, a tool not listed at all two years earlier. That

figure measures perceived prevalence, not attack frequency [2]. On June 16, 2026 the FBI issued a public service announcement on vacant-land seller impersonation built on stolen personal data, fake IDs and VoIP numbers [13].

The clearest documented U.S. deepfake attempt came in Hallandale Beach, Florida, in September 2024. A title company stopped a vacant-lot sale after the purported seller appeared on a Zoom verification call that "was clearly not a real person". The face was built from images of a missing California woman [14]. A person caught the fake by noticing non-human video and mismatched records. No detection tool was involved. The industry's standard fix, "verify by video call", has itself become an attack surface.

1.2

Listing photos carry the most weight in a buyer's decision

Listing imagery matters because buyers rely on it more than on anything else online. In NAR's 2025 Generational Trends report, 83% of buyers who searched online rated photos "very useful". Photos ranked first of 14 website features, ahead of detailed property information (79%) and floor plans (57%), and 51% of buyers found the home they bought online [15].

Most listing manipulation is "shallow": sky replacement, object removal and virtual staging, not generative deepfakes. NAR treats these edits under Code of Ethics Article 12's duty to "present a true picture", and urges agents to review all marketing images before publishing [16, 17]. The only quantitative prevalence estimate of altered U.S. listing photos the research found came from a vendor and was excluded from this report. Independent prevalence measurement is an open gap.

02

Insurance: a $308.6 billion baseline as policyholders take the pictures

2.1

The baseline and its caveats

The Coalition Against Insurance Fraud's widely used baseline puts U.S. insurance fraud at $308.6 billion a year, about $45 billion of it in property and casualty. CAIF also says fraud occurs in about 10% of P&C losses [18, 19]. Both figures need caveats. The total is a 2022 compilation, not a measurement. A Rutgers Law School report calls it "based on faulty research methodology" [20]. It also predates any AI-specific data, and no CAIF figure yet isolates manipulated media.

2.2

What claims professionals report

The best current window into manipulated media is Verisk's 2026 State of Insurance Fraud study. It surveyed 1,000 consumers and 300 claims professionals in December 2025 and January 2026. Verisk sells media-forensics tools, so treat the study as vendor research, and the figures are self-reported perceptions, not measured fraud rates. 99% of insurers said they had encountered manipulated or AI-altered documentation, and 76% said these submissions grew more sophisticated over the past year. 36% of consumers would consider altering a claim photo or document, including 55% of Gen Z. 13% consider creating a photo of damage that never happened acceptable [3].

99%

of surveyed insurers have encountered manipulated or AI-altered documentation (Verisk, vendor survey) [3]

36%

of surveyed consumers would consider altering a claim photo or document, including 55% of Gen Z [3]

60%

of insurer-required U.S. home inspections are now do-it-yourself, per vendor executives [23]

Most edits consumers called acceptable were benign, led by brightness or contrast (52%). 53% of insurers think at least half of policyholders who alter claim media do not realize it might be fraud [3]. Verisk and CAIF presented similar survey findings to the NAIC Antifraud Task Force, whose draft minutes record that 33% of respondents would be willing to use AI to manipulate a claim image [21].

2.3

Photo-first claims and do-it-yourself underwriting

Two structural shifts are widening the exposure. Photo-first claims are mainstream. CCC launched photo-based touchless auto estimating in 2021 with four national insurers, including USAA [22]. Underwriting has moved into the policyholder's hands as well. An estimated 60% of insurer-required U.S. home inspections are now do-it-yourself, and 80% of insureds choose a digital self-survey when offered. The software behind it is used by three of the top five U.S. personal-lines carriers [23]. Those figures come from vendor executives quoted by Insurance Journal. Once the insured takes the photo, image integrity becomes an underwriting input as well as a claims input.

The risk also runs in both directions. After Hurricane Ian, independent adjusters alleged that carriers had altered their damage reports, cutting estimates by up to 80% while keeping the adjusters' names on them. Florida closed the inquiry for lack of witness participation [24]. Evidence integrity therefore protects every party in a claim, not only the carrier.

The research found no public U.S. prosecution specifically for an AI-altered damage photo. U.S. cases involve fabricated invoices, staged physical damage and altered reports, charged under general fraud statutes. This likely reflects under-detection and charging practice more than absence.

03

Why metadata fails

3.1

Metadata proves nothing when it is present and is usually missing

EXIF and XMP are self-asserted, unsigned data. NIST's federal reference on synthetic content, AI 100-4, states that anyone copying or editing a file "can then manipulate or falsify the metadata in the versions they create and distribute". It adds that metadata "can similarly be stripped altogether, as it often is when files are shared (e.g., via social media platforms)" [4].

The underlying platform measurement is old. A 2013 IPTC test reported by the Library of Congress found that several popular services strip embedded metadata. Removing one field there often takes captions, location and capture date with it [25]. The research found no rigorous 2023–2026 measurement for individual platforms, and current blog claims contradict each other. A screenshot creates a new file whose only metadata comes from the screenshotting device.

Two implications follow. A timestamp or GPS tag in a claim photo establishes nothing on its own. Its absence is the normal state of shared media, not a red flag. Advice to put "date stamps" on home-inventory photos, as the California Department of Insurance recommends, produces exactly this kind of editable claim [26].

04

Why detection fails

4.1

Detectors guess, they degrade, and they can be defeated

NIST describes automated detection as "a constant cat-and-mouse game". Detectors "may only perform well on specific generators". In cross-generator tests after compression and resizing, cited studies report 61–70% accuracy in one paper and 50–62% in another [4]. The pattern holds across benchmarks.

EVIDENCE	FINDING	SOURCE
Meta Deepfake Detection Challenge (2,114 participants, 35,000+ models)	Winning model fell from 82.56% on public data to 65.18% on unseen videos	[5]
Deepfake-Eval-2024 (in-the-wild media from 88 websites)	Open-source detector AUC dropped ~50% (video), 48% (audio), 45% (image) versus older benchmarks	[27]
Ojha et al., UW–Madison (CVPR 2023)	Classifiers trained on GANs fail on newer diffusion and autoregressive generators	[28]
Carlini (Google) and Farid (UC Berkeley)	Flipping the lowest bit of each pixel cut a 0.95-AUC detector to 0.0005 AUC; black-box attack to 0.22	[6]
Hussain et al., UC San Diego (WACV 2021)	Adversarial perturbations bypass video deepfake detectors and survive compression	[29]

Table: Detector accuracy on unseen generators, in-the-wild media and adversarial attacks.

65.18%

accuracy of the winning Deepfake Detection Challenge model on unseen videos, down from 82.56% [5]

0.0005

AUC left after flipping the lowest bit of each pixel against a 0.95-AUC detector [6]

48%

human accuracy classifying StyleGAN2 faces, near chance [31]

DARPA's Semantic Forensics program concluded that "purely statistical detection methods are quickly becoming insufficient" [30]. Inpainting, which matters most for property and damage photos, is especially hard. NIST notes that with partially synthetic content "most of the content will be real" [4].

> **...purely statistical detection methods are quickly becoming insufficient**
>
> SEMANTIC FORENSICS (SEMAFOR) PROGRAM, DEFENSE ADVANCED RESEARCH PROJECTS AGENCY [30]

4.2

Human review does not close the gap

In PNAS, participants classified StyleGAN2 faces with 48% accuracy, near chance, and rated the synthetic faces 7.7% more trustworthy than real ones [31]. A 15,016-person MIT study found that crowds and the leading DFDC model were similarly accurate. When participants were shown an inaccurate model prediction, their own accuracy often dropped [32]. A Communications of the ACM study of 1,276 participants found detection "as good as a coin toss" across images, video and audio, and self-reported expertise did not help [7].

Insurers' own numbers point the same way. Only 32% are "very confident" identifying deepfakes. 38% say their tools miss too many altered images, and 35% report false positives [3]. The Casualty Actuarial Society's Actuarial Review reports that AI tools and consumer tactics "seem to continuously outpace" insurer detection [33].

4.3

False positives are the decisive failure

False positives are the decisive failure for these industries. NIST warns that labeling human-made content as AI-generated "can be extremely damaging" and recommends reporting detection rates at very low false-positive thresholds [4]. Mislabeling a legitimate claim photo means wrongly delaying or denying a claim. Debevoise warns insurers against exactly that kind of overcorrection [34]. A detector score can raise a flag. It cannot clear a photo.

Regulators also police accuracy claims. The FTC's final order against Workado, whose AI-content detector was advertised as "98 percent" accurate but tested at 53% on general content, requires "competent and reliable evidence" for any such claim [35, 36].

4.4

Invisible watermarks do not rescue detection

NIST reports that "most current watermarking schemes have been consistently found vulnerable to removal" [4]. A NeurIPS 2024 paper from UCSB, CMU and Berkeley proves that pixel-level invisible watermarks can be removed by adding noise and regenerating the image [37]. University of Maryland researchers showed that watermarks can also be spoofed onto real photos so they test as AI-generated [38]. Watermarks cover only cooperating generators, so a missing watermark proves nothing. A forged one could even be used to cast doubt on a genuine claim photo.

05

C2PA: the right cryptography, not yet the coverage

5.1

A Content Credential notarizes bytes and a signer, not truth

A C2PA Content Credential is a signed manifest of assertions about an asset. It is tied to the exact bytes by cryptographic hashes (a "hard binding"), so any tampering "will invalidate one or more hashes". It is signed with X.509 certificates and SHA2-256. A watermark or fingerprint (a "soft binding") can help find the manifest again after it has been stripped [39].

The standard is "currently on version 2.4". A Conformance Program launched on June 4, 2025 now governs which signers count as trusted [9, 40]. C2PA's guidance distinguishes "Valid", meaning cryptographically intact, from "Trusted", meaning the signer is on the C2PA Trust List [9]. NIST states the boundary plainly. A signature "cannot verify that the metadata is accurate; on a technical level, it merely serves a notarization function" [4]. C2PA agrees that provenance "cannot tell you whether the digital content is true, accurate or factual" [39].

> **"...cannot verify that the metadata is accurate; on a technical level, it merely serves a notarization function**
>
> NIST AI 100-4, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NOVEMBER 2024) [4]

5.2

Membership is broad, but shipped capture signing is thin

The institutional backing is real. The steering committee includes Adobe, Amazon, BBC, Google, Meta, Microsoft, OpenAI, Publicis, Sony, TikTok and Truepic [41], and C2PA reports more than 6,000 members and affiliates [42]. Shipping conformant code lags far behind. As of May 14, 2026, 54 products from 31 companies had passed conformance [9].

6,000+

C2PA members and affiliates [42]

54

products from 31 companies had passed C2PA conformance as of May 14, 2026 [9]

2.4

the current version of the C2PA standard [9]

On the capture side, the Pixel 10 is the reference implementation. Google says it is "the first to have Content Credentials built in across every photo created by Pixel Camera", at Assurance Level 2. Keys sit in hardware (StrongBox on Titan M2), each certificate signs exactly one image, and the CAs keep no logs [8]. Most other deployments label AI generation or AI edits rather than certifying real captures. The research found no primary source confirming native C2PA capture signing on iPhone, which leaves a large share of brokers, adjusters and policyholders without hardware-level capture credentials.

5.3

Display is uneven, and credentials get stripped

Display is uneven. YouTube's "Captured with a camera" label requires a C2PA 2.1+ capture tool and no edits to the sound or visuals [43]. LinkedIn acknowledged in 2024 that it could not yet identify and label all AI content [44]. C2PA's own guidance warns that screenshots and re-exports strip credentials, and that "No Content Credentials found" may simply mean they were removed [9].

NIST adds a privacy cost. Provenance metadata "could leak sensitive metadata", and when a credential points to an external URL, "attempts to access that URL can be tracked" [4]. WITNESS, which led C2PA's harms work, warns that location assertions can expose people who did not intend to share them [45].

The practical conclusion: an AI label is a negative signal that disappears when stripped, leaving content "unlabeled", not "real". Only signed capture provenance makes a positive claim that can be checked. Even that claim is limited to "captured by this device or app, unchanged since signing."

06

Regulation and evidence: converging on provenance without mandating a technology

6.1

California now requires a public link to the original listing photo

AB 723 (Chapter 497, Statutes of 2025; Bus. & Prof. Code §10140.8) took effect January 1, 2026. If a California licensee advertises with an image "altered through the use of photo editing software or artificial intelligence" to add, remove or change elements, the ad must carry a "reasonably conspicuous" disclosure "on or adjacent to the image". It must also link to a publicly accessible URL or QR code that "clearly identifies the original, unaltered image". Covered elements include fixtures, furniture, flooring, walls, paint, landscape and window views. Routine "lighting, sharpening, white balance, color correction, angle, straightening, cropping, exposure" adjustments are exempt [10].

The California Association of Realtors opposed the bill, and the Consumer Federation of California supported it [10]. Outside California, NorthstarMLS proposed guidance in May 2026 calling for disclosure and an unaltered "before" image alongside any AI-enhanced or virtually staged photo [46]. AB 723 assumes a trustworthy original exists but says nothing about how anyone can tell that a file is the original. A capture-time hash and verification link fill that gap. The statute does not require one, and using one does not create a compliance safe harbor.

6.2

The California AI Transparency Act and federal proposals

The California AI Transparency Act (SB 942, 2024, as amended by AB 853, Chapter 674, Statutes of 2025) became operative August 2, 2026. Its base duties require large GenAI providers to offer detection tools and embed latent disclosures in AI output, with $5,000 penalties per violation [47, 48]. From January 1, 2027, large online platforms must detect and surface standards-based provenance data. From January 1, 2028, manufacturers of capture devices sold in California must give users an option to embed a latent provenance disclosure: manufacturer, device and capture time [49].

Sources disagree on whether the enacted text bans tools built to remove provenance, so check the chaptered text before relying on that point. These duties fall on generators, platforms and device makers, not on capture apps. They make capture-time provenance a statutory expectation in California. At the federal level, the COPIED Act, which would have NIST develop provenance standards and outlaw tampering with provenance data, was reintroduced in April 2025 and has not passed [50].

6.3

Insurance AI rules govern the insurer's algorithms, not the claimant's evidence

The NAIC Model Bulletin on the Use of AI Systems by Insurers (adopted December 4, 2023) has been adopted in roughly half the states. Law-firm trackers counted 24 states in March 2025 and about 25 by 2026 [51, 52]. NY DFS Circular Letter No. 7 (2024) sets governance, fairness-testing and vendor-oversight expectations for AI and external data in underwriting and pricing [53]. Colorado's SB21-169 bars unfairly discriminatory use of external data and models [54].

All three regulate how insurers use AI. None sets a standard for authenticating policyholder-submitted media, and the research found no NAIC or state guidance on minimum authentication standards for claim photos. Fabricated images are already covered by general fraud statutes: California Penal Code §550 covers false statements in support of a claim, and New York Penal Law §176.05 covers written statements containing materially false information. Neither statute names AI [55, 56].

In practice, claims teams are left with Debevoise's advice to collect multiple angles, native original files, live video inspections and more site visits [34]. That is effectively a manual provenance request. Verisk's respondents see the same trajectory: 45% expect stricter proof-of-loss requirements within three to five years [3].

6.4

Evidence law already accepts hash-verified process records

Under FRE 901(a), authentication requires only evidence "sufficient to support a finding that the item is what the proponent claims it is". Rule 901(b)(9) allows authentication by "evidence describing a process or system and showing that it produces an accurate result" [57]. FRE 902(13) and (14), in force since December 2017, allow self-authentication by certification for records generated by an accurate electronic process and for data copies verified "by a process of digital identification", which the committee note ties to hash values [58]. A capture-time record with a content hash, device, time and stable verification link maps directly onto these routes.

AI-specific reform has stalled. Proposed FRE 707, on machine-generated evidence, went out for public comment from August 2025 to February 2026. In June 2026 the Standing Committee sent it back for further study [59]. A deepfake-specific draft Rule 901(c), which would require a challenger to first show evidence of fabrication, was never formally proposed and was tabled in May 2026 [60]. As of October 2026 the FRE contain no deepfake rule.

Contemporaneous provenance is therefore the strongest available answer to both a fabricated exhibit and a "liar's dividend" claim that genuine media is fake. Most insurance disputes, however, run through state courts, appraisal or arbitration under state evidence codes, which this research did not cover.

07

Capture-first provenance and two link types

7.1

What capture-first provenance changes

Capture-first provenance changes the question. Instead of asking "does this look fake?", it asks "can this file be traced to a specific capture, unchanged since?" Answering that does not require knowing which generator produced a fake. NIST frames provenance and detection as complementary. Provenance "can be applied to both synthetic and non-synthetic content", and checking it is "relatively straightforward". Content-based detection, by contrast, needs evaluation datasets NIST says are still hard to build [4]. Controlled capture also has a commercial track record in insurance, where carriers already rely on in-app capture for policyholder self-inspections [23].

Apex Verify applies this model at the app layer. Only captures made in Apex Camera earn proof, while camera-roll imports, screenshots, re-uploads and AI output never do. Each capture gets a content hash and a verification record held off-file. That record survives the platform stripping NIST describes, which matters because NIST notes that external provenance repositories "do not yet exist on a wide scale" [4].

7.2

A public link and a private link

Two kinds of link match two privacy needs.

01 **Permanent public verification link.** Suits media the owner chooses to publish, such as listing originals or marketing images. Permanence is what makes such a link citable, and it fits AB 723's requirement for a public URL or QR code [10].

02 **Revocable 24-hour private link.** Suits sensitive evidence such as interiors, damage, claims files and deposit disputes. It applies the access limits NIST recommends: people gain "greater control of their privacy if they have the ability to control or delete their metadata", and external lookups can otherwise be tracked [4]. It also matches the NIST Privacy Framework's goals of limiting observability and linkability [61].

Two design rules follow from the same sources. Lookup logging should be kept to a minimum. Revoking a link should end access but keep the integrity record, so later disputes can still be resolved.

7.3

Where brokerages and insurers should insert it

WORKFLOW	EXISTING DOCUMENTATION RULE	RECOMMENDED LINK	WHAT PROVENANCE DOES NOT ESTABLISH
Altered listing photos (California)	AB 723: disclosure plus public link to "original, unaltered image" [10]	Permanent public	That the property is free of defects not shown in frame
Rental move-in/move-out	Cal. Civ. Code §1950.5 (AB 2801): photos before tenancy (from July 1, 2025) and before/after repairs for deductions [62]	Private 24-hour to tenant or court; public if both agree	Who caused the damage
Desktop and hybrid appraisals	Fannie Mae B4-1.2-02: interested-party photos must be verified with a disinterested source; lender responsible for property data [63]	Private to appraiser and lender	Condition outside the frame; it supports but does not replace the appraiser's verification
DIY underwriting inspections	~60% of insurer-required home inspections now self-performed [23]	Private to carrier, via the broker	Hazards the insured chose not to photograph
First notice of loss / photo claims	Photo-first estimating mainstream [22]; counsel advises native originals [34]	Private to adjuster and SIU	That damage was not staged, or occurred on the stated loss date

WORKFLOW	EXISTING DOCUMENTATION RULE	RECOMMENDED LINK	WHAT PROVENANCE DOES NOT ESTABLISH
Pre-loss home inventory	CDI recommends video walk-through and date-stamped photos [26]	Private, kept by the insured	Ownership or value
Seller identity at closing	ALTA: 89% flag sellers reluctant to meet or talk [2]	Not applicable	Identity or liveness; use dedicated ID verification

Table: Where capture-first records fit existing documentation rules, and what they do not establish.

Brokers sit at the junction in both industries. Listing agents control originals, and insurance brokers often send inspection requests to clients and pass documentation on to carriers. The research found no source documenting brokers' legal duties or E&O exposure for passing on altered media, so that role is a practical recommendation, not an established obligation.

08

Honest limits

8.1

What capture provenance cannot prove

A capture record proves less than its users will want. It shows that a file came from a particular capture and has not changed since. It cannot show that the scene was not staged, that the damage occurred on the stated date, that the room is the insured's or that a caption is true. NIST warns that transparency tools "can create a false sense of trust" when legitimate content is reused out of context [4].

Photographing a screen produces a genuine capture of fake content (the "analog hole"), and the research found no primary U.S. study measuring such recapture attacks. App-level signing on iOS also sits at a lower assurance tier than the Pixel 10's hardware-backed keys. Google itself notes that any claim "is fundamentally dependent on the integrity of the application and the OS" [8]. NIST adds that compromised tools can "add or remove data before the signature is generated" [4].

> **"...is fundamentally dependent on the integrity of the application and the OS**
>
> E. LYNCH AND S. HANNA, GOOGLE SECURITY BLOG (SEPTEMBER 10, 2025) [8]

8.2

What Apex does not yet do

Apex-specific limits must be stated plainly. Apex does not yet embed certificate-backed, interoperable C2PA credentials in every exported file, so a carrier's or platform's C2PA validator cannot read its records today. Trust currently rests on Apex's own server and its continuity, not on a trust-listed X.509

signer. Adopting C2PA, ideally through the Conformance Program with a "Trusted" signer, belongs on the roadmap; it is not a current feature. Apex detects nothing: a file without an Apex record is unverified, not fake.

Requiring capture-first media also raises fairness and access questions. Not every claimant or tenant has the app. Debevoise's warning against delaying or denying legitimate claims applies equally to demanding provenance in ways that exclude people [34].

8.3

Gaps in the evidence base

This paper draws on U.S. government and regulatory documents, statutes and legislative records, standards bodies (NIST, C2PA), industry associations (NAR, ALTA, CAIF, NAIC, CAS), peer-reviewed or archival academic research, major law-firm analyses, and established U.S. trade or news outlets reporting named primary studies. Vendor-sponsored figures (Verisk, EXL/Truepic, CCC) are labeled as such in the text. Unverified vendor blogs, figures available only secondhand and non-U.S. case reports were excluded.

The evidence base has gaps of its own. The fraud-scale figures are victim-reported (IC3, FTC), industry perception surveys (ALTA, Verisk) or contested compilations (CAIF). None measures how often manipulated media actually appears in U.S. claims or listings. No independent benchmark of detector accuracy on insurance claim images was found.

Regulatory points not yet confirmed against chaptered text include AB 723's penalty mechanism, SB 942's provider thresholds and penalty amount, AB 853's removal-tool provision and the exact NAIC adoption count. The fall 2026 outcome of the federal evidence-rules process was not available at the time of writing. Detection retains a role as a complement for the large body of media that will never carry provenance, which is how NIST and leading forensics researchers frame it [4].

09

Conclusion

The core shift is from judging content to verifying chains of custody. Every failure documented here, including stripped EXIF, generator-specific detectors, removable watermarks and humans at chance, comes from trying to infer authenticity from a file that has already left the camera. Every rule now arriving, including AB 723's public original, California's 2028 capture-device option, FRE 902(14)'s hash-verified copies and insurers' push for native originals, assumes instead that authenticity can be established at capture and preserved afterward.

Brokerages and carriers do not need to win the detection arms race. They need to decide at which moments they will accept only media with a verifiable capture record: listing originals, move-in photos, self-inspections and first notice of loss. Everything else should be treated as testimony, not proof.

The second takeaway concerns privacy. A trustworthy media system for homes and claims cannot be fully public, because the most valuable evidence (interiors, damage, personal property, location) is also the most sensitive. Pairing permanent public links for published media with expiring, revocable private links for evidence responds to NIST's own warning that provenance lookups can be tracked.

For Apex, the credibility test over the next two years is interoperability. Once its capture records also travel as Trusted C2PA credentials, the off-file verification link becomes a resilience layer on top of an open standard instead of a proprietary substitute for one. Until then, its honest claim is narrow, and the narrowness is a strength: it can be verified.

Key answers

How can an insurer verify that a claim photo is real?

No detector can clear a photo: a score can raise a flag, and mislabeling a legitimate claim photo means wrongly delaying or denying a claim [4, 34]. The workable question is whether the file can be traced to a specific capture, unchanged since. Debevoise advises collecting multiple angles, native original files, live video inspections and more site visits, which is effectively a manual provenance request [34]. A capture-first record shared privately with the adjuster and SIU answers that question, but it does not establish that damage was not staged or occurred on the stated loss date.

Why do AI and deepfake detectors fail?

Detectors "may only perform well on specific generators", and cross-generator tests after compression and resizing report 61–70% and 50–62% accuracy [4]. The winning Deepfake Detection Challenge model fell from 82.56% on public data to 65.18% on unseen videos [5], and flipping the lowest bit of each pixel cut a 0.95-AUC detector to 0.0005 AUC [6]. DARPA concluded that "purely statistical detection methods are quickly becoming insufficient" [30].

Can people spot deepfakes by eye?

Not reliably. Participants classified StyleGAN2 faces with 48% accuracy, near chance, and rated the synthetic faces more trustworthy than real ones [31]. A study of 1,276 participants found detection "as good as a coin toss" across images, video and audio, and self-reported expertise did not help [7]. Only 32% of insurers are "very confident" identifying deepfakes [3].

Can EXIF metadata prove when a photo was taken?

No. EXIF and XMP are self-asserted, unsigned data that anyone copying or editing a file "can then manipulate or falsify", and metadata is often stripped when files are shared [4]. A timestamp or GPS tag in a claim photo establishes nothing on its own, and its absence is the normal state of shared media, not a red flag.

What is C2PA and what are Content Credentials?

A C2PA Content Credential is a signed manifest of assertions about an asset, tied to the exact bytes by cryptographic hashes and signed with X.509 certificates [39]. C2PA distinguishes "Valid", meaning cryptographically intact, from "Trusted", meaning the signer is on the C2PA Trust List [9]. A signature "merely serves a notarization function" and cannot verify that the metadata is accurate [4]. As of May 14, 2026, 54 products from 31 companies had passed conformance, and screenshots and re-exports strip credentials [9].

Does California AB 723 require the original of an edited listing photo?

Yes. Since January 1, 2026, a California licensee advertising with an image altered by photo editing software or AI to add, remove or change elements must carry a disclosure and a publicly accessible URL or QR code that "clearly identifies the original, unaltered image" [10]. Routine adjustments such as lighting, white balance, cropping and exposure are exempt. The statute does not say how anyone can tell that a file is the original; a capture-time hash and verification link fill that gap, but using one does not create a compliance safe harbor.

Are there rules for authenticating AI-altered insurance claim photos?

Not yet. The NAIC Model Bulletin, NY DFS Circular Letter No. 7 and Colorado's SB21-169 regulate how insurers use AI, and none sets a standard for authenticating policyholder-submitted media [51, 53, 54]. Fabricated images are covered by general fraud statutes such as California Penal Code §550 and New York Penal Law §176.05, neither of which names AI [55, 56]. 45% of Verisk's respondents expect stricter proof-of-loss requirements within three to five years [3].

What does capture-first provenance prove, and what does it not prove?

It shows that a file came from a particular capture and has not changed since, and checking it is "relatively straightforward" [4]. It cannot show that the scene was not staged, that the damage occurred on the stated date, that the room is the insured's or that a caption is true. Photographing a screen produces a genuine capture of fake content (the "analog hole"), and app-level signing on iOS sits at a lower assurance tier than hardware-backed keys [8].

Is Apex Verify an AI detector?

No. Apex Verify is capture-first provenance: only captures made in Apex Camera earn proof, each with a content hash and an off-file verification record, and a file without an Apex record is unverified, not fake. Its claim is narrow: this file came from an Apex Camera capture and is unchanged since. It does not yet embed certificate-backed, interoperable C2PA credentials in every exported file, so standard C2PA validators cannot read its records today, and trust currently rests on Apex's own server.

When should a broker use a public link versus a 24-hour private link?

Use a permanent public verification link for media the owner chooses to publish, such as listing originals or marketing images; it fits AB 723's requirement for a public URL or QR code [10]. Use a revocable 24-hour private link for sensitive evidence such as interiors, damage, claims files, deposit disputes and underwriting self-inspections, because provenance lookups can be tracked [4, 61]. Revoking a link should end access but keep the integrity record, so later disputes can still be resolved.

References

[1] Federal Bureau of Investigation, Internet Crime Complaint Center. Internet Crime Report 2025. https://www.ic3.gov/AnnualReport/Reports/2025_IC3Report.pdf

[2] National Mortgage News. "Rising seller impersonation fraud risk a worry for lenders" (reporting ALTA's 2026 Seller Impersonation Fraud Study), September 14, 2026. https://www.nationalmortgagenews.com/news/rising-seller-impersonation-fraud-risk-a-worry-for-lenders

[3] Claims Journal. Coverage of Verisk's 2026 State of Insurance Fraud study (vendor research), March 17, 2026. https://www.claimsjournal.com/news/national/2026/03/17/336303.htm

[4] National Institute of Standards and Technology. NIST AI 100-4: Reducing Risks Posed by Synthetic Content: An Overview of Technical Approaches to Digital Content Transparency, November 2024. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-4.pdf

[5] Meta AI. "Deepfake Detection Challenge results: An open initiative to advance AI," June 12, 2020. https://ai.meta.com/blog/deepfake-detection-challenge-results-an-open-initiative-to-advance-ai/

[6] Carlini, N. & Farid, H. "Evading Deepfake-Image Detectors with White- and Black-Box Attacks." arXiv:2004.00622, 2020. https://arxiv.org/abs/2004.00622

[7] Cooke, D., Edwards, A., Barkoff, S. & Kelly, K. "As Good As A Coin Toss: Human detection of AI-generated images, videos, audio, and audiovisual stimuli." Communications of the ACM 68 (2025); arXiv:2403.16760. https://arxiv.org/abs/2403.16760

[8] Google Security Blog (Lynch, E. & Hanna, S.). Pixel 10 and Android trusted images with C2PA Content Credentials, September 10, 2025. https://security.googleblog.com/2025/09/pixel-android-trusted-images-c2pa-content-credentials.html

[9] Coalition for Content Provenance and Authenticity (C2PA). Content Credentials Deployment Guidance, Version 1.0, July 8, 2026. https://c2pa.org/wp-content/uploads/sites/33/2026/07/Content-Credentials-Deployment-Guidance.pdf

[10] California Legislature, AB 723 (Pellerin), Chapter 497, Statutes of 2025 (Bus. & Prof. Code §10140.8), bill text and votes via CalMatters Digital Democracy. https://calmatters.digitaldemocracy.org/bills/ca_202520260ab723

[11] Federal Trade Commission. Consumer Protection Data Spotlight, "Rental scams hit home: $65 million in reported losses," December 22, 2025. https://www.ftc.gov/news-events/data-visualizations/data-spotlight/2025/12/rental-scams-hit-home-65-million-reported-losses

[12] American Land Title Association. Seller Impersonation Fraud Study (2026). https://alta.org/business-operations/research-initiatives-and-resources/critical-issue-studies/seller-impersonation-fraud-study

[13] American Land Title Association. "FBI Warns of Rising Vacant Land Seller Impersonation Scams" (summarizing the FBI IC3 PSA of June 16, 2026), June 18, 2026. https://alta.org/news-and-publications/news/20260618-FBI-Warns-of-Rising-Vacant-Land-Seller-Impersonation-Scams

[14] WPLG Local 10 (Miami). "'It shocked me': Scammers now using AI deepfakes to fraudulently sell real estate," September 19, 2024. https://local10.com/news/local/2024/09/19/it-shocked-me-scammers-now-using-ai-deepfakes-to-fraudulently-sell-real-estate

[15] National Association of REALTORS®. 2025 Home Buyers and Sellers Generational Trends Report (Exhibits 3-4 and 3-6), April 2025. https://cms.nar.realtor/sites/default/files/2025-03/2025-home-buyers-and-sellers-generational-trends-report-04-01-2025.pdf

[16] National Association of REALTORS®. Code of Ethics, Article 12 (2014 update). https://www.nar.realtor/sites/default/files/policies/2007/Article-12-Updated-2014.pdf

[17] National Association of REALTORS®, Real Estate News. "When does enhancing listing photos go too far?" https://www.nar.realtor/news/real-estate-news/when-does-enhancing-listing-photos-go-too-far

[18] PropertyCasualty360. "$308.6 billion: The real annual cost of insurance fraud in the U.S." (Coalition Against Insurance Fraud study with Colorado State University Global), August 25, 2022. https://propertycasualty360.com/2022/08/25/308-6-billion-the-real-annual-cost-of-insurance-fraud-in-the-u-s/?amp=1

[19] Coalition Against Insurance Fraud. Fraud Stats (accessed October 29, 2026). https://insurancefraud.org/fraud-stats/

[20] Feinman, J. Misinformation About Insurance Fraud. Rutgers Law School, Center for Risk and Responsibility, November 2024. https://crr.rutgers.edu/wp-content/uploads/sites/22/Misinformation-About-Insurance-Fraud-11-2024.pdf

[21] National Association of Insurance Commissioners. Antifraud (D) Task Force, draft national-meeting minutes (Verisk and CAIF presentation). https://content.naic.org/sites/default/files/national_meeting/17_Minutes-AFTF_1.pdf

[22] CCC Intelligent Solutions, via Business Wire. Launch of CCC Estimate – STP photo-based touchless estimating, October 28, 2021. https://www.businesswire.com/news/home/20211028005115/en

[23] Insurance Journal. DIY home inspections report (figures from EXL and Truepic executives), May 27, 2025. https://www.insurancejournal.com/news/national/2025/05/27/824890.htm

[24] Insurance Journal. Florida investigation of alleged altered Hurricane Ian adjuster reports, March 13, 2023. https://www.insurancejournal.com/news/southeast/2023/03/13/711901.htm

[25] Library of Congress, The Signal. Ashenfelder, M. (guest post by D. Riecks). "Social Media Networks Stripping Data from Your Digital Photos," April 11, 2013. https://blogs.loc.gov/thesignal/2013/04/social-media-networks-stripping-data-from-your-digital-photos/

[26] California Department of Insurance. Home Inventory Guide (updated June 22, 2023). https://www.insurance.ca.gov/01-consumers/105-type/95-guides/03-res/upload/Home-Inventory-Guide-Updated-062223.pdf

[27] Chandra, N. A. et al. "Deepfake-Eval-2024: A Multi-Modal In-the-Wild Benchmark of Deepfakes Circulated in 2024." arXiv:2503.02857, 2025. https://arxiv.org/abs/2503.02857

[28] Ojha, U., Li, Y. & Lee, Y. J. "Towards Universal Fake Image Detectors that Generalize Across Generative Models." arXiv:2302.10174, 2023. https://arxiv.org/abs/2302.10174

[29] Hussain, S., Neekhara, P., Jere, M., Koushanfar, F. & McAuley, J. "Adversarial Deepfakes: Evaluating Vulnerability of Deepfake Detectors to Adversarial Examples." WACV 2021; arXiv:2002.12749. https://arxiv.org/abs/2002.12749

[30] Defense Advanced Research Projects Agency. Semantic Forensics (SemaFor) program page. https://www.darpa.mil/research/programs/semantic-forensics

[31] Nightingale, S. J. & Farid, H. "AI-synthesized faces are indistinguishable from real faces and more trustworthy." PNAS 119(8), e2120481119, 2022. https://www.pnas.org/doi/10.1073/pnas.2120481119

[32] Groh, M., Epstein, Z., Firestone, C. & Picard, R. "Deepfake detection by human crowds, machines, and machine-informed crowds." PNAS 119(1), e2110013119, 2022. https://pmc.ncbi.nlm.nih.gov/articles/PMC8740705

[33] Casualty Actuarial Society, Actuarial Review. "Deepfakes put insurers in deep water," May–June 2026. https://digital.casact.org/issue/may-june-2026/deepfakes-put-insurers-in-deep-water/

[34] Debevoise & Plimpton. "Use of AI-Generated Images for Fake Insurance Claims and Other Frauds," January 20, 2026. https://www.debevoise.com/insights/publications/2026/01/use-of-ai-generated-images-for-fake-insurance

[35] Federal Trade Commission. "FTC Order Requires Workado to Back Up Artificial Intelligence Detection Claims," April 28, 2025. https://www.ftc.gov/news-events/news/press-releases/2025/04/ftc-order-requires-workado-back-artificial-intelligence-detection-claims

[36] Federal Trade Commission. Final order approved in Workado, LLC matter, August 28, 2025. https://www.ftc.gov/node/297697

[37] Zhao, X. et al. "Invisible Image Watermarks Are Provably Removable Using Generative AI." NeurIPS 2024; arXiv:2306.01953. https://arxiv.org/abs/2306.01953

[38] Saberi, M. et al. "Robustness of AI-Image Detectors: Fundamental Limits and Practical Attacks." arXiv:2310.00076, 2023 (rev. 2024). https://arxiv.org/abs/2310.00076

[39] C2PA. C2PA Explainer, Specification 2.2. https://spec.c2pa.org/specifications/specifications/2.2/explainer/Explainer.html

[40] Content Authenticity Initiative. "Raising the bar for trust: Introducing the C2PA Conformance Program," June 2025. https://contentauthenticity.org/blog/raising-the-bar-for-trust-introducing-the-c2pa-conformance-program

[41] C2PA. Membership (accessed October 2026). https://c2pa.org/membership/

[42] C2PA. "The C2PA launches Content Credentials 2.3 and celebrates 5 years of impact across the digital ecosystem" (updated February 23, 2026). https://c2pa.org/the-c2pa-launches-content-credentials-2-3-and-celebrates-5-years-of-impact-across-the-digital-ecosystem/

[43] PetaPixel. "YouTube's new Content Credentials point toward a more transparent future for video content," October 16, 2024. https://petapixel.com/2024/10/16/youtubes-new-content-credentials-point-toward-a-more-transparent-future-for-video-content

[44] Fortune. Report on LinkedIn's Content Credentials labeling challenges, November 26, 2024. https://fortune.com/2024/11/26/linkedin-content-credentials-labeling-challenges-c2pa

[45] WITNESS. "WITNESS and the C2PA Harms and Misuse Assessment Process," December 2021. https://blog.witness.org/2021/12/witness-and-the-c2pa-harms-and-misuse-assessment-process/

[46] NorthstarMLS. "Guidelines for virtual staging and AI-enhanced listing photos" (proposed), May 21, 2026. https://northstarmls.com/insights/guidelines-for-virtual-staging-and-ai-enhanced-listing-photos/

[47] California Legislature, SB 942, California AI Transparency Act, Chapter 291, Statutes of 2024, via CalMatters Digital Democracy. https://calmatters.digitaldemocracy.org/bills/ca_202320240sb942

[48] California Legislature, AB 853, Chapter 674, Statutes of 2025, bill record via Plural Policy. https://open.pluralpolicy.com/ca/bills/20252026/AB853

[49] California Senate Judiciary Committee. Bill analysis, AB 853 (2025). https://sjud.senate.ca.gov/media/1926

[50] U.S. Senate Committee on Commerce, Science, and Transportation. "Cantwell, Blackburn, Heinrich Reintroduce Bipartisan Bill to Increase Transparency, Combat AI Deepfakes" (COPIED Act), April 2025. https://www.commerce.senate.gov/index.php/2025/4/cantwell-blackburn-heinrich-reintroduce-bipartisan-bill-to-increase-transparency-combat-ai-deepfakes-put-journalists-artists-songwriters-back-in-control-of-their-content

[51] National Association of Insurance Commissioners. Implementation map, Model Bulletin on the Use of Artificial Intelligence Systems by Insurers (status as of May 1, 2026). https://content.naic.org/sites/default/files/legal-adoption-map-ai-model-bulletin.pdf

[52] Quarles & Brady. "Nearly Half of States Have Now Adopted NAIC Model Bulletin on Insurers' Use of AI," March 2025. https://quarles.com/newsroom/publications/nearly-half-of-states-have-now-adopted-naic-model-bulletin-on-insurers-use-of-ai

[53] Debevoise & Plimpton. "NYDFS Adopts Final Circular on Use of AI" (Circular Letter No. 7, July 11, 2024), July 2024. https://www.debevoise.com/insights/publications/2024/07/nydfs-adopts-final-circular-on-use-of-ai

[54] Colorado Division of Insurance. "SB21-169: Protecting Consumers from Unfair Discrimination in Insurance Practices." https://doi.colorado.gov/for-consumers/sb21-169-protecting-consumers-from-unfair-discrimination-in-insurance-practices

[55] California Penal Code §550 (via Justia). https://law.justia.com/codes/california/code-pen/part-1/title-13/chapter-5/section-550/

[56] New York Penal Law §176.05 (New York State Senate). https://www.nysenate.gov/legislation/laws/PEN/176.05

[57] Federal Rules of Evidence, Rule 901 (Legal Information Institute, Cornell Law School). https://www.law.cornell.edu/rules/fre/rule_901

[58] Federal Rules of Evidence, Rule 902 (Legal Information Institute, Cornell Law School). https://www.law.cornell.edu/rules/fre/rule_902

[59] Proskauer Rose, Minding Your Business. "Machine-Generated Evidence Challenges the Federal Rules: Inside the Contested Proposed Federal Rule of Evidence 707," July 29, 2026. https://www.proskauer.com/blog/machine-generated-evidence-challenges-the-federal-rules-inside-the-contested-proposed-federal-rule-of-evidence-707

[60] Bloomberg Law, U.S. Law Week. "Proposed AI Evidentiary Rules Punted Due to Lack of Consensus," May 2026. https://news.bloomberglaw.com/us-law-week/proposed-ai-evidentiary-rules-punted-due-to-lack-of-consensus

[61] National Institute of Standards and Technology. NIST Privacy Framework (Control function, Disassociated Processing category). https://www.nist.gov/privacy-framework

[62] California Legislature, AB 2801, Chapter 280, Statutes of 2024 (Civ. Code §1950.5), chaptered text via LegiScan. https://legiscan.com/CA/text/AB2801/2023

[63] Fannie Mae. Selling Guide B4-1.2-02, Desktop Appraisals (dated December 10, 2025). https://selling-guide.fanniemae.com/sel/b4-1.2-02/desktop-appraisals

[64] Federal Bureau of Investigation, Internet Crime Complaint Center. 2022 Internet Crime Report (real estate: 11,727 victims, $396,932,821 in losses). https://www.ic3.gov/AnnualReport/Reports/2022_IC3Report.pdf

Shot in Apex. Or it doesn't count.

Apex Verify is capture-first provenance for photos and videos: sealed at capture, shared as a permanent public link or a revocable 24-hour private link.
apexverify.app